TECHNICAL RESEARCH REPORT

A Note on an LQG Regulator with Markovian
Switching and Pathwise Average Cost

by M.K. Ghosh, A. Arapostathis and
S.I. Marcus

T.R. 94-61

ISR

INSTITUTE FOR SYSTEMS RESEARCH

Report Documentation Page

1. REPORT DATE **1994**	2. REPORT TYPE	3. DATES COVERED **00-00-1994 to 00-00-1994**

4. TITLE AND SUBTITLE **A Note on an LQG Regulator with Markovian Switching and Pathwise Average Cost**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical Engineering,Institute for Systems Research,University of Maryland,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT **see report**

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **8**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

A NOTE ON AN LQG REGULATOR WITH MARKOVIAN SWITCHING AND PATHWISE AVERAGE COST*

Mrinal K. Ghosh†, Aristotle Arapostathis‡ and Steven I. Marcus§

ABSTRACT. We study a linear system with a Markovian switching parameter perturbed by white noise. The cost function is quadratic. Under certain conditions, we find a linear feedback control which is almost surely optimal for the pathwise average cost over the infinite planning horizon.

1. Introduction. We study a parameterized linear system perturbed by white noise. The parameters are randomly switching from one state to the other and are modelled as a finite state Markov chain; the values of the parameter and the state of the linear system are assumed to be known to the controller. The objective is to minimize a quadratic cost over the infinite planning horizon. Such dynamics arise quite often in numerous applications involving systems with multiple modes or failure modes, such as fault tolerant control systems, multiple target tracking, flexible manufacturing systems, etc. [3], [4], [9].

For a finite planning horizon, the problem is well understood, but difficulties arise when the planning horizon is infinite (very large in practice) and one looks for a steady state solution. Due to constant perturbation by the white noise, the total cost is usually infinite, rendering the total cost criterion inappropriate for measuring performance. In this situation, one studies the (long-run) average cost criterion. Here we study the optimal control problem of such a system with pathwise average cost. Pathwise results are very important in practical applications, since we often deal with a single realization. This problem for a very general hybrid system has been studied in [5], where we have established the existence of an almost surely optimal stationary Markov control and have characterized it as a minimizing selector of the Hamiltonian associated with the corresponding dynamic programming equations. When specializing to the LQG case, the existence results in [5] carry through with minor modifications, but the characterization results do not, since the boundedness of the drift is crucially used to derive the dynamic programming equations. In this note we sketch the derivation of the existence result for the LQG problem from [5]. We then characterize

the optimal control by solving the dynamic programming equations via Riccati equations. Similar dynamics, but with an overtaking optimality criterion, have been recently studied in [6].

Our paper is structured as follows. Section 2 deals with the problem description. The main results are contained in Section 3. Section 4 concludes the paper with some remarks.

2. Problem Description. For the sake of notational convenience, we treat the scalar case. The higher dimensional case can be treated in an analogous manner. Let $S(t)$ be a (continuous time) Markov chain taking values in $\mathcal{S} = \{1, 2, \ldots, N\}$ with generator $\Lambda = [\lambda_{ij}]$ such that $\lambda_{ij} > 0$, $i \neq j$ (this condition can be relaxed. See Remark 3.1(v).). Let $X(t)$ be a one-dimensional process given by

$$(2.1) \qquad \begin{aligned} dX(t) &= \left[A\big(S(t)\big)X(t) + B\big(S(t)\big)u(t) \right] dt + \sigma\big(S(t)\big)W(t) \\ X(0) &= X_0 \end{aligned}$$

for $t \geq 0$, where $W(\cdot)$ is a one-dimensional standard Brownian motion independent of $S(\cdot)$ and X_0. $u(t)$ is a real valued nonanticipative process satisfying

$$(2.2) \qquad \int_0^T u^2(t) dt < \infty \quad \text{a.s. (almost surely)}$$

for each $T > 0$, $A(i)$, $B(i)$, $\sigma(i)$ are scalars such that $B(i) \neq 0$ and $\sigma(i) > 0$ for each i. We will often write A_i, B_i, etc. instead of $A(i)$, $B(i)$. The nonanticipative process $u(t)$ satisfying (2.2) is called an *admissible control*. It is called a *stationary Markov control* if $u(t) = v\big(X(t), S(t)\big)$ for a measurable map $v : \mathbb{R} \times \mathcal{S} \to \mathbb{R}$. With an abuse of notation, the map v itself is called a stationary Markov control. A stationary Markov control $v(x, i)$ is called a *linear feedback control* if $v(x, i) = k_i x$, where k_i are scalars, $i = 1, \ldots, N$.

Under a stationary Markov control v, the hybrid process $\big(X(t), S(t)\big)$ is a strong (time-homogeneous) Markov process [4]. Let v be a stationary Markov control. Under v, the point $(x, i) \in \mathbb{R} \times \mathcal{S}$ is said to be *recurrent under v* if

$$P_{x,i}^v\big(X(t_n) = x, S(t_n) = i, \text{ for a sequence } t_n \uparrow \infty\big) = 1,$$

where $P_{x,i}^v$ is the measure under v and with initial condition $X(0) = x$, $S(0) = i$, in (2.1). A point(x, i) is *transient* under v, if

$$P_{x,i}^v\big(|X(t_n)| \to \infty, \text{ as } t \to \infty\big) = 1.$$

For higher dimension$d$, the point $(x, i) \in \mathbb{R}^d \times \mathcal{S}$, $d > 1$, is said to be recurrent under v, if given any $\epsilon > 0$,

$$P_{x,i}^v\big(X(t_n) \in B(x, \epsilon), S(t_n) = i, \text{ for a sequence } t_n \uparrow \infty\big) = 1.$$

If all points (x, i) are recurrent, then the hybrid process $\big(X(t), S(t)\big)$ is called recurrent. It is shown in [5] that under our assumption, for any stationary Markov control, $\big(X(t), S(t)\big)$

is either recurrent or transient. A recurrent $(X(t), S(t))$ will admit a unique (up to a constant multiple) σ-finite invariant measure on $\mathbb{R} \times \mathcal{S}$ (resp. $\mathbb{R}^d \times \mathcal{S}$ for higher dimension). The hybrid process $(X(t), S(t))$ is called positive recurrent if it is recurrent and admits a finite invariant measure. For characterization and other details concerning recurrence and ergodicity of this class of hybrid systems we refer to [5]. A stationary Markov control v is called *stable* if the corresponding process $(X(t), S(t))$ is positive recurrent. The cost function $c(x, i, u) : \mathbb{R} \times \mathcal{S} \times \mathbb{R} \to \mathbb{R}_+$ is given by

$$(2.3) \qquad c(x, i, u) = C(i)x^2 + D(i)u^2,$$

where $C_i > 0$, $D_i > 0$ for each i. We wish to minimize a.s.

$$(2.4) \qquad \limsup_{T \to \infty} \frac{1}{T} \int_0^T \left[C(S(t))X^2(t) + D(S(t))u^2(t) \right] dt$$

over all admissible controls. Note that in (2.4) there is no expectation. We say that an admissible policy $u(\cdot)$ is a.s. optimal if there exists a constant ρ^* such that

$$(2.5) \qquad \limsup_{T \to \infty} \frac{1}{T} \int_0^T \left[C(S(t))X^2(t) + D(S(t))u^2(t) \right] dt = \rho^* \ a.s.,$$

where $X(t)$ is the solution of (2.1) under $u(\cdot)$, and for any other admissible policy $\tilde{u}(\cdot)$

$$(2.5) \qquad \limsup_{T \to \infty} \frac{1}{T} \int_0^T \left[C(S(t))\tilde{X}^2(t) + D(S(t))\tilde{u}^2(t) \right] dt = \rho^* \ a.s.,$$

where $\tilde{X}(t)$ is the solution of (2.1) under $\tilde{u}(\cdot)$, and initial condition $\tilde{X}(0) = X_0$. Note that in (2.5) and (2.6) the two measures with respect to which the 'a.s.' qualifies may be defined on different measurable spaces. Our goal is to show the existence of an a.s. optimal control and then find an a.s. optimal linear feedback control.

3. Main Results. First we prove the following result.

Lemma 3.1. *The set of stable Markov policies is nonempty.*

Proof. If each $B_i > 0$, then for an appropriate scalar K, the Markov policy $u(t) = KX(t)$ can be shown to be stable. Indeed, $w(x) = x^2$ would be a Liapunov function for the hybrid process $(X(t), S(t))$ under the control $u(t) = KX(t)$. If each B_i is not positive, we proceed as follows. Choose scalars K_i such that $A_i + K_i B_i < 0$. Consider the system of equations in Q_i

$$(3.1) \qquad 2(A_i + K_i B_i)Q_i + 1 + \sum_{j=1}^N \lambda_{ij} Q_j = 0.$$

Since $\lambda_{ii} < 0$ for each i, it is not difficult to show that (3.1) has a unique positive solution Q_i. In this case, the stationary Markov control (linear feedback)

$$u(t) = K(S(t))X(t)$$

can be shown to be stable. Indeed, $w(x,i) = Q_i x^2$ is a Liapunov function for the process $\big(X(t), S(t)\big)$ under the control $u(t) = K\big(S(t)\big)X(t)$. $\square$

For a stable stationary Markov control v, there exists a unique invariant probability measure, denoted by η_v, for the corresponding process $\big(X(t), S(t)\big)$, and

$$(3.2) \quad \rho_v := \lim_{T \to \infty} \frac{1}{T} \int_0^T \big[C\big(S(t)\big)X^2(t) + D\big(S(t)\big)v^2\big(X(t), S(t)\big)\big] dt$$

$$= \sum_{i=1}^N \int_{\mathbb{R}} \big(C_i x^2 + D_i v^2(x,i)\big) \eta_v(dx, i) \quad \text{a.s.}$$

If v, as above, is a stable linear feedback, then it can be shown as in [10] that $\rho_v < \infty$. Let

$$(3.3) \qquad\qquad\qquad \rho^* = \inf_v \rho_v$$

where the infimum is over all stable linear feedback controls. Clearly, we look for a stable linear feedback control v^* such that $\rho^* = \rho_{v^*}$ and v^* is a.s. optimal among all admissible controls. Since $C_i > 0$, the cost function c in (2.3) penalizes the unstable behavior. In other words, the cost penalizes the drift of the process away from some compact set, requiring the optimal control to exert some kind of a "centripetal force" pushing the process back towards this compact set. Thus, the optimal control gains the desired stability property. In the framework of [5], it is easily seen that the penalizing condition (A5)

$$\liminf_{|x| \to \infty} \inf_u c(x, i, u) > \rho^*$$

is satisfied. Thus by the results of [5], we have the following existence result.

Theorem 3.1. *There exists a stable stationary Markov control which is a.s. optimal.*

We now proceed to find a stable linear feedback control which is a.s. optimal. To this end, we study the dynamic programming equation given by

$$(3.4) \qquad \frac{1}{2}\sigma_i^2 V''(x,i) + \min_u \big[(A_i x + B_i u)V'(x,i)\big] + \sum_{j=1}^N \lambda_{ij} V(x,j) = \rho,$$

where ρ is a scalar and $V : \mathbb{R} \times \mathcal{S} \to \mathbb{R}$. Since the drift is unbounded, the dynamic programming treatment of [5] is not applicable here. We look for a trial solution of (3.4), of the form

$$(3.5) \qquad\qquad\qquad V(x,i) = Q_i x^2 + R_i,$$

where Q_i and R_i are to be determined. Following the usual procedure, we find that the minimizing selector in (3.4) is given by

$$(3.6) \qquad\qquad\qquad v(x,i) = -\frac{B_i Q_i x}{D_i}$$

4

where the Q_i's are the unique positive solution of the algebraic Riccati system

$$(3.7) \qquad 2A_iQ_i - \frac{B_i^2Q_i^2}{D_i} + C_i + \sum_{j=1}^{N} \lambda_{ij}Q_j = 0$$

and the R_i's are given by

$$(3.8) \qquad \sigma_i^2 Q_i + \sum_{j=1}^{N} \lambda_{ij}R_j = \rho.$$

Note that (3.8) is an underdetermined system of equations in R_i, ρ, $i = 1, 2, \ldots, N$. Also, if the R_i's satisfy (3.8), then so do $R_i + k$ for any constant k.

Lemma 3.2. *Fix an $i_0 \in S$. Then there exists a unique solution (R_i, ρ) to (3.8) satisfying $R_{i_0} = 0$.*

Proof. We have for any $T > 0$,

$$(3.9) \qquad \begin{aligned} E\,R(S(T)) - E\,R(S(0)) &= E\int_0^T \sum_{j=1}^{N} \lambda_{S(t),j}R_j \\ &= \rho T - \frac{1}{T}\int_0^T \sigma^2\big(S(t)\big)Q\big(S(t)\big)dt, \end{aligned}$$

where the second equality follows from (3.8). Dividing (3.9) by T, letting $T \to \infty$, and using the fact that the chain $S(t)$ is irreducible and ergodic, we have

$$(3.10) \qquad \begin{aligned} \rho &= \lim_{T\to\infty} \frac{1}{T} E\int_0^T \sigma^2\big(S(t)\big)Q\big(S(t)\big)dt \\ &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \sigma^2\big(S(t)\big)Q\big(S(t)\big)dt \quad \text{a.s.} \\ &= \sum_{i=1}^{N} \sigma_i^2 Q_i \pi(i) \end{aligned}$$

where $\pi = [\pi(1), \ldots, \pi(N)]'$ is the (unique) invariant probability measure of $S(t)$. Let R_i be defined by

$$(3.11) \qquad R_i = E\Big[\int_0^{\tau_{i_0}} \big(\sigma^2(S(t))Q(S(t)) - \rho\big)dt \,\big|\, S(0) = i\Big]$$

where $\tau_{i_0} = \inf\{t > 0 \mid S(t) = i_0\}$. Then, using Dynkin's formula, it is seen that R_i satisfies (3.8) with $R_{i_0} = 0$. Let (R_i', ρ') be another solution to (3.8) such that $R_{i_0}' = 0$. As before, we have

$$\rho' = \sum_{i=1}^{N} \sigma_i^2 Q_i \pi(i) = \rho.$$

Let $M(t) = R\big(S(t)\big) - R'\big(S(t)\big)$. Then using (3.8), $M(t)$ is easily seen to be a martingale which converges a.s. Since $S(t)$ is positive recurrent, it visits every state infinitely often. Hence $R_i - R_i'$ must be a constant. Thus $R_i - R_i' = R_{i_0} - R_{i_0}' = 0$. $\quad\square$

In view of these results, using Ito's formula and the pathwise analysis of [5], the following result is now apparent.

Theorem 3.2. *For the LQG problem (2.1)–(2.3), the linear feedback control given by (3.6) is a.s. optimal, where the Q_i's are determined by (3.8). The pathwise optimal cost is given by (3.9).*

Some comments are in order now.

Remark 3.1.

(i) The condition $B_i \neq 0$ for each i can be relaxed. If for some i, $B_i = 0$, then the above result will still hold if (A_i, B_i) are stochastically stabilizable in a certain sense [7], [9].

(ii) For the multidimensional case, if B_i, C_i, σ_i are positive definite, then all the above results will hold. But the positive definiteness of B_i is a very strong condition, since in many cases the B_i's may not even be square matrices. In such a case if we assume that (A_i, B_i) are stochastically stabilizable, then the above results will again hold. Sufficient conditions for stochastic stabilizability are given in [2], [7], [9]. If C_i is not positive definite, then the cost does not necessarily penalize the unstable behavior, as discussed in the foregoing. Thus the condition (A5) of [5] is not satisfied. In this case under a further detectability condition, the optimality can be obtained in a restricted class of stationary Markov controls [2].

(iii) Let ρ be as in (3.9). Then for any admissible policy $u(t)$ it can be shown by the pathwise analysis in [5] that

$$\liminf_{T \to \infty} \frac{1}{T} \int_0^T \big[C\big(S(t)\big) X^2(t) + D\big(S(t)\big) u^2(t) \big] dt \geq \rho \quad \text{a.s.}$$

This establishes the optimality of the linear feedback control v (3.6) in a much stronger sense, viz. the most "pessimistic" pathwise average cost under v is no worse than the most "optimistic" pathwise average cost under any admissible control.

(iv) For $T > 0$, let $V(x, i, T)$ denote the optimal expected cost for the finite horizon $[0, T]$. Then it can be shown as in [2] that

$$\lim_{T \to \infty} \frac{1}{T} V(x, i, T) = \rho$$

where ρ is as in (3.9). Thus the finite horizon value function approaches the optimal pathwise average cost as the length of the horizon increases to infinity. Thus for large T, the linear feedback control (3.6) would be a reasonably good nearly optimal control for the finite horizon case. This would be particularly useful in practical applications since it is computationally more economical to solve the algebraic Riccati system than the Riccati system of differential equations.

(v) The condition $\lambda_{ij} > 0$ can be relaxed to the condition that the chain $S(t)$ is irreducible (and hence ergodic). The existence part in [5] can be suitably modified to make the necessary claim here. In the dynamic programming part, the existence of a unique solution in Lemma 3.2 is clearly true under the irreducibility condition.

4. Conclusions. In this note, we have studied the pathwise optimality of an LQG regulator with Markovian switching parameters. We have assumed that the Markovian parameters are known to the controllers. This is an ideal situation. In practice the controllers may not have a complete knowledge of these parameters. In this case, one usually studies the corresponding minimum variance filter. Unfortunately, this filter is almost always infinite dimensional [9]. A computationally efficient suboptimal filter has been developed in [1], [8]. We hope that our results will be useful in the dual control problem arising in this situation.

REFERENCES

1. H. A. P. Blom, *An efficient filter for abruptly changing systems*, Proc. 23rd IEEE Conf. Decision and Control, Las Vegas (1984), 656–658.
2. M. D. Fragoso and E. M. Hemerly, *Optimal control for a class of noisy linear systems with Markovian jumping parameters and quadratic cost*, Int. J. Syst. Sci. **22** (1991), 2553–2561.
3. M. K. Ghosh, A. Arapostathis and S. I. Marcus, *An optimal control problem arising in flexible manufacturing systems*, Proc. 30th IEEE Conf. Decision and Control, Brighton, England (1991), 1844–1849.
4. ———, *Optimal control of switching diffusions with application to flexible manufacturing systems*, SIAM J. Control Optim. **31** (1993), 1183–1204.
5. ———, *Ergodic control of switching diffusions*, preprint.
6. A. Haurie and A. Leizarowitz, *Overtaking optimal regulation and tracking of piecewise diffusion linear systems*, SIAM J. Control Optim. **30** (1992), 816–837.
7. Y. Ji and H. J. Chizeck, *Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control*, IEEE Trans. Automat. Contr. **AC-35** (1990), 777–788.
8. S.I. Marcus and E.K. Westwood, *On asymptotic approximation for some nonlinear filtering problems*, Proc. IFAC Triennial Congress, Budapest, Hungary, July 4-6, 1984, Vol. VII, 36–41.
9. M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, New York, 1990.
10. W. M. Wonham, *Random differential equations in control theory*, Probabilistic Methods in Applied Mathematics (A. T. Bharucha-Reid, ed.), vol. 2, Academic Press, New York, 1970, pp. 131–212.